

KARDAR-PARISI-ZHANG EQUATION AND LARGE DEVIATIONS FOR RANDOM WALKS IN WEAK RANDOM ENVIRONMENTS

IVAN CORWIN AND YU GU

ABSTRACT. We consider the transition probabilities for random walks in $1 + 1$ dimensional space-time random environments (RWRE). For critically tuned weak disorder we prove a sharp large deviation result: after appropriate rescaling, the transition probabilities for the RWRE evaluated in the large deviation regime, converge to the solution to the stochastic heat equation (SHE) with multiplicative noise (the logarithm of which is the KPZ equation). We apply this to the exactly solvable Beta RWRE and additionally present a formal derivation of the convergence of certain moment formulas for that model to those for the SHE.

1. INTRODUCTION

We consider random walks in $1 + 1$ dimensional space-time random environments (RWRE). The environment is specified by a sequence of zero-mean i.i.d. random variables (whose probability measure is denoted \mathbb{P} and expectation operator \mathbb{E}):

$$\omega = \{\omega_{i,j} : i \in \mathbb{Z}_{\geq 0}, j \in \mathbb{Z}\},$$

and a parameter $\varepsilon > 0$. For each realization of ω , and each ε we consider a measure $\mathbf{P}^{\omega,\varepsilon}$ on one-dimensional nearest neighbor walks $\{S_n\}_{n \geq 0}$ started at the origin ($S_0 = 0$) which jump up or down according to the probabilities

$$(1.1) \quad \mathbf{P}^{\omega,\varepsilon}(S_{n+1} = S_n + 1) = \frac{1 + \varepsilon^{\frac{1}{2}} \omega_{n,S_n}}{2}, \quad \mathbf{P}^{\omega,\varepsilon}(S_{n+1} = S_n - 1) = \frac{1 - \varepsilon^{\frac{1}{2}} \omega_{n,S_n}}{2}.$$

We assume that the ω and ε are such that all probabilities lie in $[0, 1]$. When $\varepsilon = 0$, this becomes the measure of a simple symmetric random walk (SSRW), and we distinguish this measure by writing it as \mathbf{P}^0 (we use $\mathbf{E}^{\omega,\varepsilon}$ and \mathbf{E}^0 to denote the respective expectation operators). The parameter ε allows us to tune the strength of the disorder around that of the SSRW. In this paper we will consider the transition probability $\mathbf{P}^{\omega,\varepsilon}(S_N = y)$ as a random variable (it inherits the randomness of ω). We are interested in the large deviation regime whereby $y = vN$ for some $v \in (0, 1)$.

For ω and ε fixed, subject to some hypotheses on ω , [10] proves that $N^{-1} \log \mathbf{P}^{\omega,\varepsilon}(S_N = vN)$ has a limit (which is called the rate function). That rate function is written in terms of the Legendre transform of another $N \rightarrow \infty$ limit. In the special case when the random variables $\frac{1}{2}(1 + \varepsilon^{\frac{1}{2}} \omega_{n,S_n})$ are distributed according to the Beta(α, β) distribution, [4] proves an explicit and simple formula for the rate function. Further, [4] proves that after centering $\log \mathbf{P}^{\omega,\varepsilon}(S_N \geq vN)$ by N times the rate function, the result fluctuates like $N^{1/3}$ times a GUE Tracy-Widom random variable. (This result was only proved for certain ranges of parameters α, β, v though should hold in general away from law of large numbers velocity.) Note that the result of [4] involved the tail probabilities instead of the transition probabilities. A similar result should hold in both cases – see the recent non-rigorous physics work of [9] regarding the transition probability fluctuations. The occurrence of cube-root fluctuations and the GUE Tracy-Widom distribution demonstrates a relation between the Beta RWRE and the KPZ universality class [7].

To further elucidate the connection between the RWRE and KPZ universality class, and to motivate our main results, let us observe how the RWRE transition probability is equal to the partition function for a directed polymer model. (We utilize the notational conventions from [4] to make comparison with those results easier.) For $N \in \mathbb{Z}_{\geq 0}$ and $x \in \mathbb{Z}_{>0}$, we consider paths starting from $(0, 1)$ and ending at (N, x) , which are allowed to make right steps by adding e_1 and diagonal steps by adding $e_1 + e_2$ (here $e_1 = (1, 0)$ and $e_2 = (0, 1)$ are standard basis vectors). The weight of a path is the product of the weights of each edge along the path, and the weight of the edge e is

defined by

$$(1.2) \quad \begin{cases} B_{i,j} & \text{if } e \text{ is the horizontal edge } (i-1, j) \rightarrow (i, j), \\ 1 - B_{i,j} & \text{if } e \text{ is the diagonal edge } (i-1, j-1) \rightarrow (i, j), \end{cases}$$

where $\{B_{i,j}\}$ is a sequence of i.i.d. random variables supported on $[0, 1]$. The point-to-point polymer partition function $Z(N, x)$ then equals the sum over all paths from $(0, 1)$ to (N, x) of the above defined path weights, and it satisfies the recursion relation

$$(1.3) \quad Z(N, x) = Z(N-1, x)B_{N,x} + Z(N-1, x-1)(1 - B_{N,x})$$

with initial data $Z(0, x) = \mathbf{1}\{x = 1\}$. This polymer has the special property that weights leading into a particular vertex always sum to 1. Let us note that the above recursion for $Z(N, x)$ is a special case of a random averaging process (RAP) which satisfies

$$Z(N, x) = \sum_{k: |k| \leq M} Z(N-1, x+k)B_{N,x}(k),$$

where $B_{i,j} = \{B_{i,j}(k) : -M \leq k \leq M, \sum_{k: |k| \leq M} B_{i,j}(k) = 1\}$ are i.i.d. probability vectors. One may try to generalize some of the results we prove herein to the RAP, though we do not pursue that here.

The RWRE transition probability is related to this partition function via a time reversal. If we let

$$(1.4) \quad \mathbf{P}^{\omega, \varepsilon}(S_{n+1} = S_n + 1) = B_{n, S_n}, \quad \mathbf{P}^{\omega, \varepsilon}(S_{n+1} = S_n - 1) = 1 - B_{n, S_n},$$

(i.e., $\omega_{i,j} = 2\varepsilon^{-\frac{1}{2}}(B_{i,j} - \frac{1}{2})$) then for fixed N and x we have the equality in law

$$(1.5) \quad Z(N, x) = \mathbf{P}^{\omega, \varepsilon}(S_N = N - 2x + 2).$$

A similar result holds for the RAP whereby the resulting RWRE may have jump sizes randomly sampled from $\{-M, \dots, M\}$ according to the environment; see [3].

Directed polymer partition functions (in fact their logarithms) are expected to show KPZ-class fluctuations under general choices of weights. This conjecture is far from proved, having only been demonstrated for certain exactly solvable models – see the review [8] for various recent references. The aforementioned result for the Beta RWRE (or equivalently polymer) thus fits into this conjecture.

For directed polymer models with i.i.d. weights on vertices (instead of edges) of the form $e^{\beta\omega}$, [1] introduced *intermediate disorder* or *weak noise* scaling whereby as N goes to infinity, x scales like $N^{1/2}$ and β like $N^{-1/4}$. Under that scaling, the partition function converges to a continuum partition function whereby a Brownian bridge moves through a space-time white noise potential and assigns a weight to each path given by the exponential of the integral of the white noise along the path. That partition function is equal, via the Feynman-Kac representation, to the solution to the stochastic heat equation (SHE) with multiplicative white noise

$$\partial_t \mathcal{Z}(t, x) = \frac{1}{2} \partial_{xx} \mathcal{Z}(t, x) + \mathcal{Z}(t, x) \dot{W}(t, x)$$

with delta initial data $\mathcal{Z}(0, x) = \delta_x$. See, for example, [1] and references therein regarding the definition of this equation and white noise \dot{W} . For reference, note that the logarithm of the stochastic heat equation solves the KPZ equation, and the weak noise scaling is such that the KPZ equation remains invariant under it.

The results of [1] was proved via the convergence of the discrete chaos series for the polymer partition function to that of the continuum Wiener chaos series for $\mathcal{Z}(t, x)$. Since we will make use of this chaos series, let us recall it here for delta initial data:

$$(1.6) \quad \mathcal{Z}(t, x) = p(t, x) + \sum_{k=1}^{\infty} \int_{\Delta_k(t) \times \mathbb{R}^k} \prod_{\ell=1}^{k+1} p(t_\ell - t_{\ell-1}, x_\ell - x_{\ell-1}) \prod_{\ell=1}^k W(dt_\ell, dx_\ell),$$

with

$$(1.7) \quad \Delta_k(t) = \{(t_1, \dots, t_k) : 0 < t_1 < \dots < t_k < t\},$$

$t_0 = 0$, $t_{k+1} = t$, $x_0 = 0$, $x_{k+1} = x$ and the standard heat kernel $p(t, x) = (2\pi t)^{-1/2} \exp\{-x^2/2t\}$. See [1] for the definition of the multiple stochastic integrals against space-time white noise.

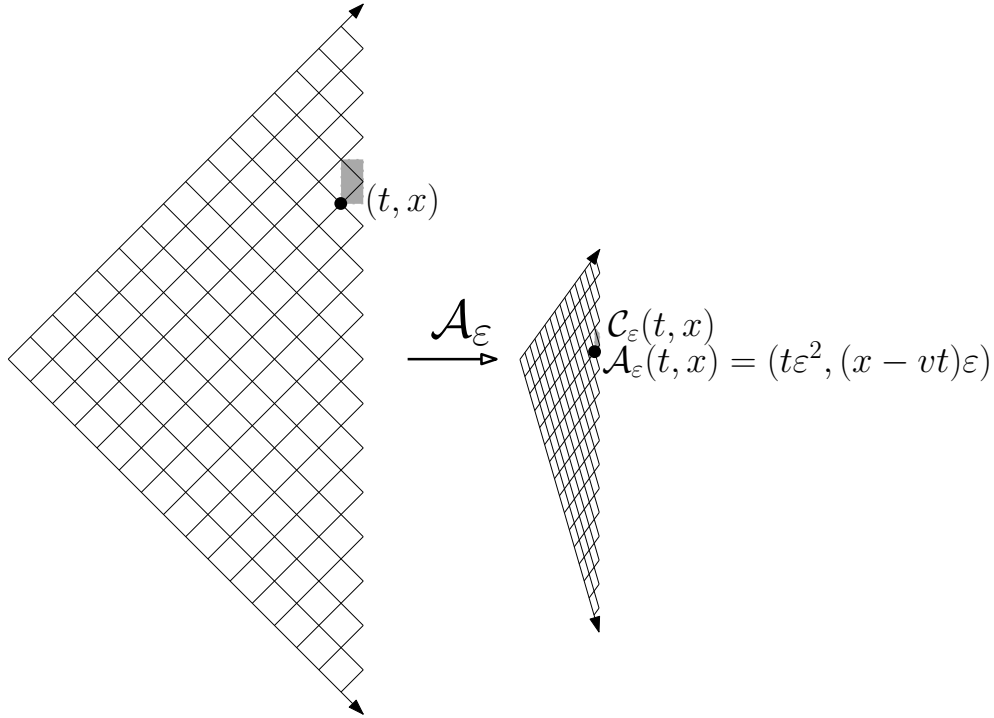


FIGURE 1. On the left is the subspace of \mathbb{Z}_2^2 on which the RWRE trajectory lives. The image under the map \mathcal{A}_ε is shown on the right. The grey box is the cell $\mathcal{C}_\varepsilon(t, x)$ indexed by (t, x) .

1.1. Main result. Inspired by the intermediate disorder regime of the polymer model explored in [1], we seek here to analyze $\mathbf{P}^{\omega, \varepsilon}(S_N = y)$ when $N \sim t/\varepsilon^2$ and $y \sim vt/\varepsilon^2 + x/\varepsilon$ for some fixed speed $v \in (0, 1)$ and $t > 0, x \in \mathbb{R}$. Notice that this scaling corresponds with that used in the polymer setup since the noise is tuned by a factor of $\varepsilon^{1/2}$ around its deterministic value $1/2$. In light of this similarity, we expect that after appropriate scaling, $\mathbf{P}^{\omega, \varepsilon}(S_N = y)$ should converge to the solution (perhaps up to simple scaling of coordinates) of the SHE with delta initial data. That is exactly what we prove here.

Before stating our theorem, we introduce notation which identifies points in $\mathbb{R}_+ \times \mathbb{R}$ with nearby points in the ε -scaled lattice on which the random walk lives. This notation will also be used in the proofs.

The RWRE paths lie along the nonnegative time, even sublattice of \mathbb{Z}^2 which we denote by

$$\mathbb{Z}_2^2 = \{(t, x) : t \in \mathbb{Z}_{\geq 0}, t + x \text{ is even}\}.$$

For v fixed and $\varepsilon > 0$ define the affine transformation $\mathcal{A}_\varepsilon : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ by

$$\mathcal{A}_\varepsilon(t, x) = (t\varepsilon^2, (x - vt)\varepsilon),$$

and denote the image of \mathbb{Z}_2^2 under \mathcal{A}_ε by \mathbb{T}_ε . For $(t, x) \in \mathbb{Z}_2^2$ let $\mathcal{C}_\varepsilon(t, x)$ be the image under \mathcal{A}_ε of the rectangle with vertices $(t, x), (t + 1, x), (t + 1, x + 2), (t, x + 2)$, which is a parallelogram. For convention we assume that the bottom and left edges of the rectangle are included whereas the top and right edges are not. As $(t, x) \in \mathbb{Z}_2^2$ varies, the $\mathcal{C}_\varepsilon(t, x)$ form a disjoint partition of $\mathbb{R}_+ \times \mathbb{R}$ into cells indexed by the bottom left point (t, x) – we refer to \mathcal{C}_ε as a tessellation and the $\mathcal{C}_\varepsilon(t, x)$ as its cells. See Figure 1 for an illustration of these definitions. For any function which is defined on \mathbb{Z}_2^2 (e.g. the transition probability) we may extend its domain to $\mathbb{R}_+ \times \mathbb{R}$ by the convention that the value on each cell is constant and equal to the value at its bottom left corner. Also, for any (t, x) , we define $(t_\varepsilon, x_\varepsilon)$ to be the unique point in \mathbb{T}_ε which lies in the same cell as (t, x) . By these conventions, we have that the probability $\mathbf{P}^{\omega, \varepsilon}$ in the left-hand side of (1.8) remains unchanged as (t, x) varying within the cell associated with $(t_\varepsilon, x_\varepsilon)$.

We are now prepared to state our main theorem.

Theorem 1.1. Fix any distribution \mathbb{P} on ω so that for ε small, the probabilities in (1.1) are in $[0, 1]$ and let $\sigma = \sqrt{2\mathbb{E}\{\omega_{i,j}^2\}}$. Then, for fixed $t > 0, x \in \mathbb{R}$ and $v \in (0, 1)$, we have

$$(1.8) \quad \frac{1}{\varepsilon} e^{\frac{t\varepsilon}{\varepsilon^2} \mathcal{I}(v) + \frac{x\varepsilon}{\varepsilon} \mathcal{I}'(v)} \mathbf{P}^{\omega, \varepsilon} \left(S_{\frac{t\varepsilon}{\varepsilon^2}} = v \frac{t\varepsilon}{\varepsilon^2} + \frac{x\varepsilon}{\varepsilon} \right) \Rightarrow \mathcal{U}(t, x)$$

in distribution (with respect to the measure \mathbb{P} on the ω) as $\varepsilon \rightarrow 0$, where $\mathcal{U}(t, x)$ solves the following SHE (with scaled coefficients)

$$(1.9) \quad \partial_t \mathcal{U}(t, x) = \frac{1}{2} (1 - v^2) \partial_{xx} \mathcal{U}(t, x) + v \sigma \mathcal{U}(t, x) \dot{W}(t, x), \quad \mathcal{U}(0, x) = 2\delta(x),$$

and

$$(1.10) \quad \mathcal{I}(v) = \frac{1-v}{2} \log(1-v) + \frac{1+v}{2} \log(1+v).$$

Theorem 1.1 implies that

$$\log \mathbf{P}^{\omega, \varepsilon} \left(S_{\frac{t\varepsilon}{\varepsilon^2}} = v \frac{t\varepsilon}{\varepsilon^2} + \frac{x\varepsilon}{\varepsilon} \right) + \frac{t\varepsilon}{\varepsilon^2} \mathcal{I}(v) + \frac{x\varepsilon}{\varepsilon} \mathcal{I}'(v) - \log \varepsilon \Rightarrow \log \mathcal{U}(t, x)$$

in distribution, i.e., the recentered logarithm of the transition probability converges in distribution to the Hopf-Cole solution to the KPZ equation with narrow wedge initial data [2].

Remark 1.2. For a SSRW starting from the origin (i.e., the $\varepsilon = 0$ case of the RWRE), a slight refinement of Cramer's theorem (see, for example, Lemma A.1) shows that

$$\frac{1}{\varepsilon} e^{\frac{t\varepsilon}{\varepsilon^2} \mathcal{I}(v) + \frac{x\varepsilon}{\varepsilon} \mathcal{I}'(v)} \mathbf{P}^0 \left(S_{\frac{t\varepsilon}{\varepsilon^2}} = v \frac{t\varepsilon}{\varepsilon^2} + \frac{x\varepsilon}{\varepsilon} \right) \rightarrow U(t, x)$$

as $\varepsilon \rightarrow 0$, where U solves the heat equation (with scaled coefficient)

$$(1.11) \quad \partial_t U(t, x) = \frac{1}{2} (1 - v^2) \partial_{xx} U(t, x), \quad U(0, x) = 2\delta(x).$$

Comparing (1.9) with (1.11), we observe that the effect of the weak random environment on the transition probability is manifested as a multiplicative space-time white noise in the limiting equation.

In light of the connection (1.5) between the RWRE and directed polymers, the proof of Theorem 1.1 also implies a similar convergence result for the polymer partition function $Z(N, x)$. The Beta polymer [4] corresponds with taking $B_{i,j}$ i.i.d. as $\text{Beta}(\alpha, \beta)$ random variables with $\alpha, \beta > 0$. By the relation $\omega_{i,j} = 2\varepsilon^{-\frac{1}{2}}(B_{i,j} - \frac{1}{2})$, fixing the distribution of the $B_{i,j}$ corresponds with taking $\omega_{i,j}$ dependent upon ε . If we tune the parameters of the $\text{Beta}(\alpha, \beta)$ random variables so that $\alpha = \beta = \varepsilon^{-1}$, then the resulting $\omega_{i,j}$ has a non-trivial $\varepsilon \rightarrow 0$ limit and the same method of proof as for Theorem 1.1 applies and yields the following.

Theorem 1.3. For the Beta polymer, fix $t > 0, x \in \mathbb{R}$ and $\gamma \in (0, \frac{1}{2})$. Let $\alpha = \beta = \varepsilon^{-1}$, we have

$$\frac{1}{\varepsilon} e^{\frac{t\varepsilon}{\varepsilon^2} \mathcal{I}(1-2\gamma) - \frac{2x\varepsilon}{\varepsilon} \mathcal{I}'(1-2\gamma)} Z \left(\frac{t\varepsilon}{\varepsilon^2}, \gamma \frac{t\varepsilon}{\varepsilon^2} + \frac{x\varepsilon}{\varepsilon} \right) \Rightarrow \mathcal{V}(t, x)$$

in distribution as $\varepsilon \rightarrow 0$, where $\mathcal{V}(t, x)$ solves

$$(1.12) \quad \partial_t \mathcal{V}(t, x) = \frac{1}{2} \gamma (1 - \gamma) \partial_{xx} \mathcal{V}(t, x) + \frac{1-2\gamma}{\sqrt{2}} \mathcal{V}(t, x) \dot{W}(t, x), \quad \mathcal{V}(0, x) = \frac{\gamma}{1-\gamma} \delta(x).$$

Remark 1.4. By (1.5), we have

$$Z \left(\frac{t\varepsilon}{\varepsilon^2}, \gamma \frac{t\varepsilon}{\varepsilon^2} + \frac{x\varepsilon}{\varepsilon} \right) = \mathbf{P}^{\omega, \varepsilon} \left(S_{\frac{t\varepsilon}{\varepsilon^2}} = (1-2\gamma) \frac{t\varepsilon}{\varepsilon^2} - \frac{2x\varepsilon}{\varepsilon} + 2 \right),$$

thus the tessellation, and consequently the meaning of $(t_\varepsilon, x_\varepsilon)$, in Theorem 1.3 should be defined with $v = 1 - 2\gamma$.

As an independent confirmation of this result, in Section 3 we demonstrate how known formulas for moments of the Beta polymer partition function converge in the above scaling to those of the limiting SHE. Note that our weak convergence result does not imply convergence of moments (though it is reasonable to expect that such a stronger form of convergence may hold).

Let us briefly sketch our approach to prove Theorem 1.1. The transition probability $\mathbf{P}^{\omega, \varepsilon}(S_N = y)$ is the sum of the probabilities of all directed paths connecting $(0, 0)$ to (N, y) in \mathbb{Z}_2^2 . These probabilities are products of the terms on the right-hand side of (1.1) and can be expanded into powers of ε . Indexing these sums in terms of the degree of the power of ε yields what is called a “polynomial chaos” in the ω noise; see Lemma 2.1. The solution to the SHE also admits an expansion in term of Wiener chaos (1.6) and the proof then reduces to showing convergence of the polynomial to Wiener chaos series. For this, we apply the framework developed in [6] (which generalizes the results for polymers studied in [1]) whereby a general criteria is given for such a convergence to hold. The key criteria to confirm is the L^2 convergence of the (deterministic) coefficients of each polynomial chaos to the corresponding coefficients of the Wiener chaos; see Lemma 2.3. Since the path of the RWRE coincides with that of the SSRW, the coefficients appearing in our setting are just joint transition probabilities of the SSRW and can be computed and bounded explicitly; see the sharp large deviation result for the SSRW presented in Appendix A.

1.2. Acknowledgements. I.C. was partially supported by the NSF through DMS-1208998, the Clay Mathematics Institute through a Clay Research Fellowship, the Institute Henri Poincaré through the Poincaré Chair, and the Packard Foundation through a Packard Fellowship for Science and Engineering. Y.G. was partially supported by the NSF through DMS-1613301. We would like to thank the anonymous referees for several helpful suggestions.

2. PROOF OF THEOREMS 1.1 AND 1.3: CONVERGENCE OF THE POLYNOMIAL CHAOS

The proof of Theorem 1.1 (and similarly of Theorem 1.3) follows from the following two lemmas. After stating these lemmas, we complete the proof of the theorems and then devote the rest of this section to the proof of the lemmas. As in the introduction, we use S_n to denote the trajectory of a nearest neighbor walk on \mathbb{Z} , $\mathbf{P}^{\omega, \varepsilon}$ to denote the RWRE measure on such a walk given ω and ε , and \mathbf{P}^0 to denote the SSRW measure. Below, we also abbreviate points $(i, j) \in \mathbb{Z}_2^2$ by $z = (i, j)$, or $z_l = (i_l, j_l)$ when referring to multiple such points.

The following lemma expresses the random transition probability $\mathbf{P}^{\omega, \varepsilon}(S_N = y)$ in terms of a polynomial chaos series with respect to the i.i.d. random variables ω . It is easily proved by expanding the product formula for the transition probability in terms of powers of ε .

Lemma 2.1. *For any $(N, y) \in \mathbb{Z}_2^2$, we have*

$$(2.1) \quad \mathbf{P}^{\omega, \varepsilon}(S_N = y) = \mathbf{P}^0(S_N = y) + \sum_{k=1}^N \varepsilon^{\frac{k}{2}} \sum_{(z_1, \dots, z_k)} \psi_k(z_1, \dots, z_k) \omega_{z_1} \dots \omega_{z_k},$$

where the summation $\sum_{(z_1, \dots, z_k)}$ is over all possible $0 \leq i_1 < \dots < i_k \leq N-1$ and $j_1, \dots, j_k \in \mathbb{Z}$, and the expansion coefficients are given by

$$(2.2) \quad \begin{aligned} \psi_k(z_1, \dots, z_k) &= \frac{1}{2^k} \mathbf{P}^0(S_{i_1} = j_1) \\ &\quad \times [\mathbf{P}^0(S_{i_2-i_1-1} = j_2 - j_1 - 1) - \mathbf{P}^0(S_{i_2-i_1-1} = j_2 - j_1 + 1)] \\ &\quad \times [\mathbf{P}^0(S_{i_3-i_2-1} = j_3 - j_2 - 1) - \mathbf{P}^0(S_{i_3-i_2-1} = j_3 - j_2 + 1)] \\ &\quad \times \dots \\ &\quad \times [\mathbf{P}^0(S_{N-i_k-1} = y - j_k - 1) - \mathbf{P}^0(S_{N-i_k-1} = y - j_k + 1)]. \end{aligned}$$

For $t > 0$ and $x \in \mathbb{R}$ recall the conventions introduced before Theorem 1.1 through which we associated (t, x) with a pair $(t_\varepsilon, x_\varepsilon) \in \mathbb{T}_\varepsilon$. We will study the RWRE transition probability when $N = t_\varepsilon/\varepsilon^2$ and $y = vt_\varepsilon/\varepsilon^2 + x_\varepsilon/\varepsilon$. Towards this end, let us rewrite the chaos series for $\mathbf{P}^{\omega, \varepsilon}(S_N = y)$ in terms of rescaled coordinates. For $z = (i, j) \in \mathbb{Z}_2^2$, denote $z_\varepsilon = \mathcal{A}_\varepsilon(i, j) \in \mathbb{T}_\varepsilon$ and for a sequence of $z_{\varepsilon, 1}, \dots, z_{\varepsilon, k} \in \mathbb{T}_\varepsilon$ define

$$(2.3) \quad \psi_{\varepsilon, k}(z_{\varepsilon, 1}, \dots, z_{\varepsilon, k}) = \psi_k(\mathcal{A}_\varepsilon^{-1} z_{\varepsilon, 1}, \dots, \mathcal{A}_\varepsilon^{-1} z_{\varepsilon, k})$$

with ψ_k given by (2.2). Also define a sequence of i.i.d. random variables indexed by $z_\varepsilon \in \mathbb{T}_\varepsilon$ by $\tilde{\omega}_{z_\varepsilon} = \omega_{\mathcal{A}_\varepsilon^{-1}z_\varepsilon}$. With these rescaled coordinates, we have the following polynomial chaos series.

Corollary 2.2. *For fixed $t > 0, x \in \mathbb{R}$, we have*

$$(2.4) \quad \mathbf{P}^{\omega, \varepsilon} \left(S_{\frac{t_\varepsilon}{\varepsilon^2}} = v \frac{t_\varepsilon}{\varepsilon^2} + \frac{x_\varepsilon}{\varepsilon} \right) = \mathbf{P}^0 \left(S_{\frac{t_\varepsilon}{\varepsilon^2}} = v \frac{t_\varepsilon}{\varepsilon^2} + \frac{x_\varepsilon}{\varepsilon} \right) + \sum_{k=1}^{\frac{t_\varepsilon}{\varepsilon^2}} \varepsilon^{\frac{k}{2}} \sum_{(z_{\varepsilon,1}, \dots, z_{\varepsilon,k}) \in \mathbb{T}_\varepsilon^k} \psi_{\varepsilon,k}(z_{\varepsilon,1}, \dots, z_{\varepsilon,k}) \tilde{\omega}_{z_{\varepsilon,1}} \dots \tilde{\omega}_{z_{\varepsilon,k}},$$

with the summation $\sum_{(z_{\varepsilon,1}, \dots, z_{\varepsilon,k}) \in \mathbb{T}_\varepsilon^k}$ restricted to $0 \leq i_{\varepsilon,1} < \dots < i_{\varepsilon,k} \leq \frac{t_\varepsilon}{\varepsilon^2} - 1$.

It remains now to prove that as ε goes to zero, the above series converges to the SHE chaos series (after some minor rescaling of coordinates) given in (1.6). Owing to the general machinery given in [6, Section 2.3], the main technical challenge in achieving this convergence is to prove the L^2 -convergence of $\psi_{\varepsilon,k}(z_{\varepsilon,1}, \dots, z_{\varepsilon,k})$ to the corresponding SHE chaos series coefficients. To state this result, we introduce the L^2 -space for this convergence. Let $\Delta_k(t)$ be the time simplex defined in (1.7). As per our conventions, the domain of the function $\psi_{\varepsilon,k}$ extends to all of $\Delta_k(t) \times \mathbb{R}^k$ by replacing each (t_l, x_l) by the corresponding $(t_{\varepsilon,l}, x_{\varepsilon,l})$. If for two distinct choices of l , the resulting $t_{\varepsilon,l}$ coincide, then set $\psi_{\varepsilon,k}$ equal to zero. The function $\psi_{\varepsilon,k}$ may also be extended to unordered times $t_1, \dots, t_k \in \mathbb{R}_+$ by setting it equal to the corresponding value for the ordered times.

Let $p_{\sigma^2}(t, x) = (2\pi\sigma^2 t)^{1/2} \exp\{-x^2/2\sigma^2 t\}$ be the density of normal distribution $N(0, \sigma^2 t)$.

Lemma 2.3. *For fixed $t > 0, x \in \mathbb{R}, k \geq 1$, let $t_0 = 0, x_0 = 0, t_{k+1} = t, x_{k+1} = x$. Then*

$$(2.5) \quad \frac{1}{\varepsilon^{1+k}} e^{\frac{t_\varepsilon}{\varepsilon^2} \mathcal{I}(v) + \frac{x_\varepsilon}{\varepsilon} \mathcal{I}'(v)} \psi_{\varepsilon,k}((t_1, x_1), \dots, (t_k, x_k)) \rightarrow 2(2v)^k \prod_{l=1}^{k+1} p_{1-v^2}(t_l - t_{l-1}, x_l - x_{l-1})$$

in $L^2(\Delta_k(t) \times \mathbb{R}^k)$. In addition,

$$(2.6) \quad \lim_{M \rightarrow \infty} \limsup_{\varepsilon \rightarrow 0} \sum_{k > M} \left\| \frac{1}{\varepsilon^{1+k}} e^{\frac{t_\varepsilon}{\varepsilon^2} \mathcal{I}(v) + \frac{x_\varepsilon}{\varepsilon} \mathcal{I}'(v)} \psi_{\varepsilon,k} \right\|_{L^2(\Delta_k(t) \times \mathbb{R}^k)}^2 = 0.$$

Now we can prove the main result.

Proof of Theorem 1.1. First observe that the left-hand side of (1.8) is equal to $\frac{1}{\varepsilon} e^{\frac{t_\varepsilon}{\varepsilon^2} \mathcal{I}(v) + \frac{x_\varepsilon}{\varepsilon} \mathcal{I}'(v)}$ times the polynomial chaos on the right-hand side of (2.4). We apply the general criteria given in [6, Theorem 2.3] to prove convergence of a polynomial chaos to a Wiener chaos. There are three criteria which we must confirm to apply these results:

- (i) The random variables $\tilde{\omega}$ are i.i.d. with $\mathbb{E}\{\tilde{\omega}\} = 0$ and $\mathbb{E}\{\tilde{\omega}^2\} < \infty$.
- (ii) The volume of the cells in the tessellation \mathcal{C}_ε is $2\varepsilon^3$. By Lemma A.1 and Lemma 2.3, we have,

$$\frac{1}{\varepsilon} e^{\frac{t_\varepsilon}{\varepsilon^2} \mathcal{I}(v) + \frac{x_\varepsilon}{\varepsilon} \mathcal{I}'(v)} \mathbf{P}^0 \left(S_{\frac{t_\varepsilon}{\varepsilon^2}} = v \frac{t_\varepsilon}{\varepsilon^2} + \frac{x_\varepsilon}{\varepsilon} \right) \rightarrow 2p_{1-v^2}(t, x)$$

and, in $L^2(\Delta_k(t) \times \mathbb{R}^k)$,

$$\begin{aligned} & \frac{1}{\varepsilon} e^{\frac{t_\varepsilon}{\varepsilon^2} \mathcal{I}(v) + \frac{x_\varepsilon}{\varepsilon} \mathcal{I}'(v)} \varepsilon^{\frac{k}{2}} \psi_{\varepsilon,k}((t_1, x_1), \dots, (t_k, x_k)) |\mathcal{C}_\varepsilon|^{-\frac{k}{2}} \\ &= \frac{1}{2^{\frac{k}{2}}} \frac{1}{\varepsilon^{1+k}} e^{\frac{t_\varepsilon}{\varepsilon^2} \mathcal{I}(v) + \frac{x_\varepsilon}{\varepsilon} \mathcal{I}'(v)} \psi_{\varepsilon,k}((t_1, x_1), \dots, (t_k, x_k)) \\ &\rightarrow 2(\sqrt{2}v)^k \prod_{l=1}^{k+1} p_{1-v^2}(t_l - t_{l-1}, x_l - x_{l-1}). \end{aligned}$$

- (iii) By Lemma 2.3, the tail

$$\lim_{M \rightarrow \infty} \limsup_{\varepsilon \rightarrow 0} \sum_{k > M} \sum_{(z_{\varepsilon,1}, \dots, z_{\varepsilon,k}) \in \mathbb{T}_\varepsilon^k} \left| \frac{1}{\varepsilon} e^{\frac{t_\varepsilon}{\varepsilon^2} \mathcal{I}(v) + \frac{x_\varepsilon}{\varepsilon} \mathcal{I}'(v)} \varepsilon^{\frac{k}{2}} \psi_{\varepsilon,k}(z_{\varepsilon,1}, \dots, z_{\varepsilon,k}) \right|^2 = 0.$$

It then follows from [6, Theorem 2.3] that

$$\begin{aligned} & \frac{1}{\varepsilon} e^{\frac{t_\varepsilon}{2} \mathcal{I}(v) + \frac{x_\varepsilon}{\varepsilon} \mathcal{I}'(v)} \mathbf{P}^{\omega, \varepsilon} \left(S_{\frac{t_\varepsilon}{2}} = v \frac{t_\varepsilon}{\varepsilon^2} + \frac{x_\varepsilon}{\varepsilon} \right) \\ & \Rightarrow 2p_{1-v^2}(t, x) + 2 \sum_{k=1}^{\infty} \int_{\Delta_k(t) \times \mathbb{R}^k} \prod_{l=1}^{k+1} p_{1-v^2}(t_l - t_{l-1}, x_l - x_{l-1}) \prod_{l=1}^k v \sigma W(dt_l, dx_l) \end{aligned}$$

in distribution, with $\sigma = \sqrt{2\mathbb{E}\{\omega_{i,j}^2\}}$ and $W(dt, dx)$ a space-time white noise. To complete the proof, we only need to note that the limit is the series of Wiener chaos expansion of (1.9). \square

Proof of Theorem 1.3. The results of Theorem 1.1 hold in slightly more generality than stated. In particular, the distribution (denoted \mathbb{P} and \mathbb{E}) of the ω may depend on ε (denoted \mathbb{P}_ε and \mathbb{E}_ε) so long as $\mathbb{E}_\varepsilon\{\omega_{i,j}\} = 0$ and $2\mathbb{E}_\varepsilon\{\omega_{i,j}^2\} \rightarrow \sigma^2$ as $\varepsilon \rightarrow 0$. Then, an inspection of [6, Theorem 2.3] reveals that the same conclusion holds as in Theorem 1.1. With this in mind, define ε -dependent $\omega_{i,j} = 2\varepsilon^{-1/2} (B_{i,j} - \frac{1}{2})$ with $B_{i,j} \sim \text{Beta}(\varepsilon^{-1}, \varepsilon^{-1})$, and note that $\mathbb{E}_\varepsilon\{\omega_{i,j}\} = 0$ and $\mathbb{E}_\varepsilon\{\omega_{i,j}^2\} \rightarrow \frac{1}{2}$ as $\varepsilon \rightarrow 0$. Thus, recalling that $Z(N, y) = \mathbf{P}^{\omega, \varepsilon}(S_N = N - 2y + 2)$ and applying the above noted generalization of Theorem 1.1, we achieve the conclusion of Theorem 1.3. \square

2.1. Proofs of Lemma 2.1 and Lemma 2.3. The proof of Lemma 2.1 is a straightforward calculation using the properties of the SSRW. To prove Lemma 2.3, it is clear by (2.2) and (2.3) that we need to analyze the sharp large deviation of the SSRW, i.e., the asymptotic behavior of

$$(2.7) \quad \mathcal{P}_\varepsilon(t_\varepsilon, x_\varepsilon, m_1, m_2) := \mathbf{P}^0 \left(S_{\frac{t_\varepsilon}{2} + m_1} = v \frac{t_\varepsilon}{\varepsilon^2} + \frac{x_\varepsilon}{\varepsilon} + m_2 \right)$$

for any $t > 0, x \in \mathbb{R}$ and $(m_1, m_2) \in \{(0, 0), (-1, -1), (-1, 1)\}$. This is left to Appendix A.

Proof of Lemma 2.1. By the definition of RWRE it follows that

$$(2.8) \quad \mathbf{P}^{\omega, \varepsilon}(S_N = y) = \mathbf{E}^0 \left\{ \prod_{i=0}^{N-1} \left[1 + (S_{i+1} - S_i) \varepsilon^{\frac{1}{2}} \omega_{i, S_i} \right] \mathbf{1}\{S_N = y\} \right\},$$

where \mathbf{E}^0 is with respect to the SSRW measure on S . Expanding the right-hand side of (2.8) we obtain (note below that the index of S in S_{i_l+1} is $i_l + 1$)

$$\mathbf{P}^{\omega, \varepsilon}(S_N = y) = \mathbf{P}^0(S_N = y) + \sum_{k=1}^N \varepsilon^{\frac{k}{2}} \sum_{0 \leq i_1 < \dots < i_k \leq N-1} \mathbf{E}^0 \left\{ \prod_{l=1}^k (S_{i_l+1} - S_{i_l}) \omega_{i_l, S_{i_l}} \mathbf{1}\{S_N = y\} \right\}.$$

The expectation can be evaluated as

$$\mathbf{E}^0 \left\{ \prod_{l=1}^k (S_{i_l+1} - S_{i_l}) \omega_{i_l, S_{i_l}} \mathbf{1}\{S_N = y\} \right\} = \sum_{(z_1, \dots, z_k)} \psi_k(z_1, \dots, z_k) \omega_{z_1} \dots \omega_{z_k},$$

where the summation $\sum_{(z_1, \dots, z_k)}$ is over all possible $0 \leq i_1 < \dots < i_k \leq N-1$ and $j_1, \dots, j_k \in \mathbb{Z}$, and the coefficients

$$(2.9) \quad \psi_k(z_1, \dots, z_k) = \sum_{\tau_1, \dots, \tau_k = \pm 1} \left(\prod_{l=1}^k \tau_l \right) \mathbf{P}^0 \left(\bigcap_{l=1}^k \{S_{i_l} = j_l, S_{i_l+1} = j_l + \tau_l\} \cap \{S_N = y\} \right).$$

Using the independence of increments of the SSRW in (2.9), we arrive at the expression in (2.8) and thus complete the proof. \square

Proof of Lemma 2.3. We first prove the convergence of (2.5) for fixed $0 < t_1 < \dots < t_k < t$ and $x_1, \dots, x_k \in \mathbb{R}$, then we prove the convergence is also in $L^2(\Delta_k(t) \times \mathbb{R}^k)$ for fixed k . In the end, we derive a bound on $\|\varepsilon^{-k-1} e^{\frac{t_\varepsilon}{2} \mathcal{I}(v)} e^{\frac{x_\varepsilon}{\varepsilon} \mathcal{I}'(v)} \psi_{\varepsilon, k}\|_{L^2(\Delta_k(t) \times \mathbb{R}^k)}$ for large k to prove (2.6).

For $0 < t_1 < \dots < t_k < t$ and $x_1, \dots, x_k \in \mathbb{R}$, by (2.2), (2.3), and (2.7), we have

$$(2.10) \quad \frac{1}{\varepsilon^{1+k}} e^{\frac{t_\varepsilon}{2} \mathcal{I}(v) + \frac{x_\varepsilon}{\varepsilon} \mathcal{I}'(v)} \psi_{\varepsilon, k}((t_1, x_1), \dots, (t_k, x_k)) = \frac{1}{2^k} \mathcal{P}_\varepsilon(t_1, x_1) \prod_{l=2}^{k+1} \mathcal{P}_\varepsilon(t_l, x_l, t_{l-1}, x_{l-1}),$$

where

$$\mathcal{P}_\varepsilon(t_1, x_1) = \frac{1}{\varepsilon} e^{\frac{t_{\varepsilon,1}}{2} \mathcal{I}(v) + \frac{x_{\varepsilon,1}}{\varepsilon} \mathcal{I}'(v)} \mathcal{P}_\varepsilon(t_{\varepsilon,1}, x_{\varepsilon,1}, 0, 0),$$

and for $2 \leq l \leq k+1$,

$$\begin{aligned} \mathcal{P}_\varepsilon(t_l, x_l, t_{l-1}, x_{l-1}) &= \frac{1}{\varepsilon} e^{\frac{t_\varepsilon, l - t_\varepsilon, l-1}{\varepsilon^2} \mathcal{I}(v) + \frac{x_\varepsilon, l - x_\varepsilon, l-1}{\varepsilon} \mathcal{I}'(v)} \\ &\times [\mathcal{P}_\varepsilon(t_\varepsilon, l - t_\varepsilon, l-1, x_\varepsilon, l - x_\varepsilon, l-1, -1, -1) - \mathcal{P}_\varepsilon(t_\varepsilon, l - t_\varepsilon, l-1, x_\varepsilon, l - x_\varepsilon, l-1, -1, 1)]. \end{aligned}$$

By Lemma A.1, we have $\mathcal{P}_\varepsilon(t_1, x_1) \rightarrow 2p_{1-v^2}(t_1, x_1)$, and for $2 \leq l \leq k+1$ we have $\mathcal{P}_\varepsilon(t_l, x_l, t_{l-1}, x_{l-1}) \rightarrow 4vp_{1-v^2}(t_l - t_{l-1}, x_l - x_{l-1})$. This implies the pointwise convergence

$$(2.11) \quad \frac{1}{\varepsilon^{k+1}} e^{\frac{t_\varepsilon}{\varepsilon^2} \mathcal{I}(v)} e^{\frac{x_\varepsilon}{\varepsilon} \mathcal{I}'(v)} \psi_{\varepsilon, k}((t_1, x_1), \dots, (t_k, x_k)) \rightarrow 2(2v)^k \prod_{l=1}^{k+1} p_{1-v^2}(t_l - t_{l-1}, x_l - x_{l-1}).$$

To prove the convergence in $L^2(\Delta_k(t) \times \mathbb{R}^k)$, we recall the uniform bound (A.2) which holds for large enough values of C :

$$\begin{aligned} \mathcal{P}_\varepsilon(t_1, x_1) &\leq C \left(\frac{1}{\varepsilon} \mathbf{1}\{t_1 \leq 100\varepsilon^2, |x_1| \leq C\varepsilon\} + \mathbf{1}\{t_1 > 100\varepsilon^2\} \frac{1}{\sqrt{t_1}} e^{-\frac{x_1^2}{Ct_1}} \right), \\ (2.12) \quad \mathcal{P}_\varepsilon(t_l, x_l, t_{l-1}, x_{l-1}) &\leq C \left(\frac{1}{\varepsilon} \mathbf{1}\{|t_l - t_{l-1}| \leq 100\varepsilon^2, |x_l - x_{l-1}| \leq C\varepsilon\} \right. \\ &\quad \left. + \mathbf{1}\{|t_l - t_{l-1}| > 100\varepsilon^2\} \frac{1}{\sqrt{t_l - t_{l-1}}} e^{-\frac{(x_l - x_{l-1})^2}{C(t_l - t_{l-1})}} \right). \end{aligned}$$

Letting

$$\Delta_{k, \varepsilon}(t) = \{(t_1, \dots, t_k) \in \Delta_k(t) : \min_{l=1, \dots, k+1} |t_l - t_{l-1}| \leq 100\varepsilon^2\},$$

we will show that the integral in $\Delta_{k, \varepsilon}(t) \times \mathbb{R}^k$ is negligible as $\varepsilon \rightarrow 0$. Before doing so, observe that in $\Delta_k(t) \setminus \Delta_{k, \varepsilon}(t)$, we have

$$\frac{1}{\varepsilon^{1+k}} e^{\frac{t_\varepsilon}{\varepsilon^2} \mathcal{I}(v) + \frac{x_\varepsilon}{\varepsilon} \mathcal{I}'(v)} \psi_{\varepsilon, k}((t_1, x_1), \dots, (t_k, x_k)) \leq C^k \prod_{l=1}^{k+1} \frac{1}{\sqrt{t_l - t_{l-1}}} e^{-\frac{(x_l - x_{l-1})^2}{C(t_l - t_{l-1})}}.$$

The right-hand side above is in $L^2(\Delta_k(t) \times \mathbb{R}^k)$. Thus, by dominated convergence and the pointwise convergence in (2.11) the following convergence holds in $L^2(\Delta_k(t) \times \mathbb{R}^k)$:

$$(2.13) \quad \frac{1}{\varepsilon^{1+k}} e^{\frac{t_\varepsilon}{\varepsilon^2} \mathcal{I}(v) + \frac{x_\varepsilon}{\varepsilon} \mathcal{I}'(v)} \psi_{\varepsilon, k}((t_1, x_1), \dots, (t_k, x_k)) \mathbf{1}\{\Delta_k(t) \setminus \Delta_{k, \varepsilon}(t)\} \rightarrow 2(2v)^k \prod_{l=1}^{k+1} p_{1-v^2}(t_l - t_{l-1}, x_l - x_{l-1}).$$

In addition (for our later proof of (2.6)) we have

$$\begin{aligned} &\left\| \frac{1}{\varepsilon^{1+k}} e^{\frac{t_\varepsilon}{\varepsilon^2} \mathcal{I}(v) + \frac{x_\varepsilon}{\varepsilon} \mathcal{I}'(v)} \psi_{\varepsilon, k} \mathbf{1}\{\Delta_k(t) \setminus \Delta_{k, \varepsilon}(t)\} \right\|_{L^2(\Delta_k(t) \times \mathbb{R}^k)}^2 \\ &\leq C^k \int_{\Delta_k(t) \times \mathbb{R}^k} \prod_{l=1}^{k+1} \frac{1}{t_l - t_{l-1}} e^{-\frac{(x_l - x_{l-1})^2}{C(t_l - t_{l-1})}} dt dx \\ (2.14) \quad &\leq \frac{C^k}{\sqrt{t}} e^{-\frac{x^2}{Ct}} \int_{\Delta_k(t)} \prod_{l=1}^{k+1} \frac{1}{\sqrt{t_l - t_{l-1}}} dt \\ &= \frac{C^k}{\sqrt{t}} e^{-\frac{x^2}{Ct}} t^{\frac{k-1}{2}} \int_{\Delta_k(1)} \frac{1}{\sqrt{t_1}} \frac{1}{\sqrt{t_2 - t_1}} \dots \frac{1}{\sqrt{1 - t_k}} dt \\ &\leq C^k e^{-\frac{x^2}{Ct}} t^{\frac{k}{2}-1} e^{-\frac{k \log k}{C}}, \end{aligned}$$

where the last step comes from the evaluation of the Dirichlet integral. It is clear that

$$(2.15) \quad \sum_{k=1}^{\infty} C^k e^{-\frac{x^2}{Ct}} t^{\frac{k}{2}-1} e^{-\frac{k \log k}{C}} < \infty.$$

Now we consider the integral in $\Delta_{k,\varepsilon}(t)$ with the aim of showing that it is negligible as $\varepsilon \rightarrow 0$. We change variables $t_l - t_{l-1} \mapsto \tau_l$, $x_l - x_{l-1} \mapsto y_l$ and use (2.12) to derive that

$$\begin{aligned} & \left\| \frac{1}{\varepsilon^{1+k}} e^{\frac{t_\varepsilon}{\varepsilon^2} \mathcal{I}(v) + \frac{x_\varepsilon}{\varepsilon} \mathcal{I}'(v)} \psi_{\varepsilon,k} \right\|_{L^2(\Delta_{k,\varepsilon}(t) \times \mathbb{R}^k)}^2 \\ & \leq C^k \int_{[0,\infty)^k} \int_{\mathbb{R}^k} \prod_{l=1}^{k+1} \left[\frac{1}{\varepsilon^2} \mathbf{1}\{|\tau_l| \leq 100\varepsilon^2, |y_l| \leq C\varepsilon\} + \frac{1}{\tau_l} e^{-\frac{y_l^2}{C\tau_l}} \mathbf{1}\{|\tau_l| > 100\varepsilon^2\} \right] \\ & \quad \mathbf{1}\left\{ \sum_{l=1}^{k+1} \tau_l = t, \min_l \tau_l \leq 100\varepsilon^2 \right\} \mathbf{1}\left\{ \sum_{l=1}^{k+1} y_l = x \right\} d\tau dy \\ & = C^k \sum_{\substack{A \subseteq \{1,\dots,k+1\} \\ |A| \geq 1}} I_A, \end{aligned}$$

where

$$\begin{aligned} I_A = & \int_{[0,\infty)^k} \int_{\mathbb{R}^k} \prod_{l \in A} \frac{1}{\varepsilon^2} \mathbf{1}\{|\tau_l| \leq 100\varepsilon^2, |y_l| \leq C\varepsilon\} \prod_{l \in A^c} \frac{1}{\tau_l} e^{-\frac{y_l^2}{C\tau_l}} \mathbf{1}\{|\tau_l| > 100\varepsilon^2\} \\ & \mathbf{1}\left\{ \sum_{l=1}^{k+1} \tau_l = t \right\} \mathbf{1}\left\{ \sum_{l=1}^{k+1} y_l = x \right\} d\tau dy. \end{aligned}$$

Here A^c is the complement of A in the set $\{1, \dots, k+1\}$.

We seek to control the behavior as $\varepsilon \rightarrow 0$ of all of these I_A expressions. We consider the following two cases (as well as some subcases). Throughout, all bounds are assumed to hold for ε small enough, and the constants C may change between lines.

Case 1: $A^c = \emptyset$. We have

$$I_A = \int_{[0,\infty)^k} \int_{\mathbb{R}^k} \prod_{l=1}^k \frac{1}{\varepsilon^2} \mathbf{1}\{|\tau_l| \leq 100\varepsilon^2, |y_l| \leq C\varepsilon\} \frac{1}{\varepsilon^2} \mathbf{1}\left\{ t - \sum_{l=1}^k \tau_l \leq 100\varepsilon^2, \left| x - \sum_{l=1}^k y_l \right| \leq C\varepsilon \right\} d\tau dy.$$

For fixed k , the above integral equals to zero when ε is small enough to make $\sum_{l=1}^{k+1} 100\varepsilon^2 < t$. For arbitrary k , we have

$$(2.16) \quad I_A \leq C^k \varepsilon^{k-2} \mathbf{1}\{k \geq 3\}.$$

Case 2: $A^c \neq \emptyset$. Fix some $l^* \in A^c$ and define $\tilde{A}^c = A^c \setminus \{l^*\}$. We have

$$\begin{aligned} I_A = & \int_{[0,\infty)^k} \int_{\mathbb{R}^k} \prod_{l \in A} \frac{1}{\varepsilon^2} \mathbf{1}\{|\tau_l| \leq 100\varepsilon^2, |y_l| \leq C\varepsilon\} \prod_{l \in \tilde{A}^c} \frac{1}{\tau_l} e^{-\frac{y_l^2}{C\tau_l}} \mathbf{1}\{|\tau_l| > 100\varepsilon^2\} \\ & \frac{1}{\tau_{l^*}} e^{-\frac{y_{l^*}^2}{C\tau_{l^*}}} \mathbf{1}\{|\tau_{l^*}| > 100\varepsilon^2\} \mathbf{1}\left\{ \sum_{l=1}^{k+1} \tau_l = t \right\} \mathbf{1}\left\{ \sum_{l=1}^{k+1} y_l = x \right\} d\tau dy. \end{aligned}$$

By symmetry, we change variables and assume $l^* = k+1$, so $A \cup \tilde{A}^c = \{1, \dots, k\}$, and

$$\begin{aligned} I_A = & \int_{[0,\infty)^k} \int_{\mathbb{R}^k} \prod_{l \in A} \frac{1}{\varepsilon^2} \mathbf{1}\{|\tau_l| \leq 100\varepsilon^2, |y_l| \leq C\varepsilon\} \prod_{l \in \tilde{A}^c} \frac{1}{\tau_l} e^{-\frac{y_l^2}{C\tau_l}} \mathbf{1}\{|\tau_l| > 100\varepsilon^2\} \\ & \frac{1}{t - \sum_{l=1}^k \tau_l} e^{-\frac{(x - \sum_{l=1}^k y_l)^2}{C(t - \sum_{l=1}^k \tau_l)}} \mathbf{1}\left\{ \sum_{l=1}^k \tau_l < t \right\} d\tau dy. \end{aligned}$$

Integrating all y_l for $l \in \tilde{A}^c$ yields

$$\begin{aligned} (2.17) \quad I_A \leq & C^{|\tilde{A}^c|} \int_{[0,\infty)^k} \int_{\mathbb{R}^{|A|}} \prod_{l \in A} \frac{1}{\varepsilon^2} \mathbf{1}\{|\tau_l| \leq 100\varepsilon^2, |y_l| \leq C\varepsilon\} \prod_{l \in \tilde{A}^c} \frac{1}{\sqrt{\tau_l}} \mathbf{1}\{|\tau_l| > 100\varepsilon^2\} \\ & \frac{1}{\sqrt{t - \sum_{l=1}^k \tau_l}} \frac{1}{\sqrt{t - \sum_{l \in A} \tau_l}} e^{-\frac{(x - \sum_{l \in A} y_l)^2}{C(t - \sum_{l \in A} \tau_l)}} \mathbf{1}\left\{ \sum_{l=1}^k \tau_l < t \right\} d\tau dy. \end{aligned}$$

If $\tilde{A}^c \neq \emptyset$, we integrate τ_l for $l \in \tilde{A}^c$ and follow (2.14) to obtain

$$(2.18) \quad I_A \leq C^{|\tilde{A}^c|} e^{-\frac{|\tilde{A}^c| \log |\tilde{A}^c|}{C}} \int_{[0, \infty)^{|A|}} \int_{\mathbb{R}^{|A|}} \prod_{l \in A} \frac{1}{\varepsilon^2} \mathbf{1}\{|\tau_l| \leq 100\varepsilon^2, |y_l| \leq C\varepsilon\} \left(t - \sum_{l \in A} \tau_l\right)^{\frac{|\tilde{A}^c|}{2}-1} \mathbf{1}\left\{\sum_{l \in A} \tau_l < t\right\} d\tau dy.$$

If $\tilde{A}^c = \emptyset$, then $A = \{1, \dots, k\}$ and we integrate any y_i in (2.17) to obtain

$$(2.19) \quad I_A \leq \int_{[0, \infty)^k} \int_{\mathbb{R}^{k-1}} \frac{1}{\varepsilon^2} \mathbf{1}\{|\tau_i| \leq 100\varepsilon^2\} \prod_{l \neq i, l=1}^k \frac{1}{\varepsilon^2} \mathbf{1}\{|\tau_l| \leq 100\varepsilon^2, |y_l| \leq C\varepsilon\} \frac{1}{\sqrt{t - \sum_{l=1}^k \tau_l}} \mathbf{1}\left\{\sum_{l=1}^k \tau_l < t\right\} d\tau dy.$$

We consider the following subcases of case 2.

Case 2.i: $k \leq 3$. We can choose ε small in (2.18) and (2.19) so that when $\tilde{A}^c \neq \emptyset$

$$I_A \leq C^k e^{-\frac{|\tilde{A}^c| \log |\tilde{A}^c|}{C}} t^{\frac{|\tilde{A}^c|}{2}-1} \varepsilon^{|A|} \leq C^k t^{\frac{|\tilde{A}^c|}{2}-1} \varepsilon^{|A|}.$$

and when $\tilde{A}^c = \emptyset$,

$$I_A \leq C^k t^{-1} \varepsilon^{|A|},$$

Thus in both cases we have

$$(2.20) \quad I_A \leq C^k t^{\frac{|\tilde{A}^c|}{2}-1} \varepsilon^{|A|}.$$

Case 2.ii: $k > 3$ and $|\tilde{A}^c| \geq 2$. We consider (2.18) and the same discussion as in case 2.i leads to

$$(2.21) \quad I_A \leq C^k e^{-\frac{|\tilde{A}^c| \log |\tilde{A}^c|}{C}} t^{\frac{|\tilde{A}^c|}{2}-1} \varepsilon^{|A|}.$$

Case 2.iii: $k > 3$ and $|\tilde{A}^c| = 1$. We integrate any τ_i in (2.18) to find

$$I_A \leq C^{|\tilde{A}^c|} e^{-\frac{|\tilde{A}^c| \log |\tilde{A}^c|}{C}} \int_{[0, \infty)^{|A|-1}} \int_{\mathbb{R}^{|A|}} \frac{1}{\varepsilon^2} \mathbf{1}\{|y_i| \leq C\varepsilon\} \prod_{l \neq i, l \in A} \frac{1}{\varepsilon^2} \mathbf{1}\{|\tau_l| \leq 100\varepsilon^2, |y_l| \leq C\varepsilon\} \left(t - \sum_{l \neq i, l \in A} \tau_l\right)^{1/2} \mathbf{1}\left\{\sum_{l \neq i, l \in A} \tau_l < t\right\} d\tau dy,$$

thus,

$$(2.22) \quad I_A \leq C^k e^{-\frac{|\tilde{A}^c| \log |\tilde{A}^c|}{C}} \varepsilon^{|A|-2} \sqrt{t} = C^k \sqrt{t} \varepsilon^{k-3}.$$

Case 2.iv: $k > 3$ and $\tilde{A}^c = \emptyset$. We integrate any τ_i in (2.19) to find

$$I_A \leq \int_{[0, \infty)^{k-1}} \int_{\mathbb{R}^{k-1}} \frac{1}{\varepsilon^2} \prod_{l \neq i, l=1}^k \frac{1}{\varepsilon^2} \mathbf{1}\{|\tau_l| \leq 100\varepsilon^2, |y_l| \leq C\varepsilon\} \left(t - \sum_{l \neq i, l=1}^k \tau_l\right)^{1/2} \mathbf{1}\left\{\sum_{l \neq i, l=1}^k \tau_l < t\right\} d\tau dy,$$

thus

$$(2.23) \quad I_A \leq C^k \sqrt{t} \varepsilon^{k-3}.$$

To summarize, by (2.16), (2.20), (2.21), (2.22) and (2.23), the following estimate holds for all k :

$$(2.24) \quad \left\| \frac{1}{\varepsilon^{1+k}} e^{\frac{t}{\varepsilon^2} \mathcal{I}(v) + \frac{x\varepsilon}{\varepsilon} \mathcal{I}'(v)} \psi_{\varepsilon, k} \right\|_{L^2(\Delta_{k, \varepsilon}(t) \times \mathbb{R}^k)}^2 \leq C^k \left(\varepsilon^{k-2} \mathbf{1}\{k \geq 3\} + \sum_{a=1}^k \binom{k+1}{a} \mathcal{E}(k, a) \right),$$

where $a^c = k+1-a$, $\tilde{a}^c = a^c - 1 = k-a$ (so that $a = |A|$, $a^c = |A^c|$ and $\tilde{a}^c = |\tilde{A}^c|$), and

$$\mathcal{E}(k, a) = \mathbf{1}\{k \leq 3\} t^{\frac{\tilde{a}^c}{2}-1} \varepsilon^a + \mathbf{1}\{k > 3, \tilde{a}^c \geq 2\} e^{-\frac{\tilde{a}^c \log \tilde{a}^c}{C}} t^{\frac{\tilde{a}^c}{2}-1} \varepsilon^a + \mathbf{1}\{k > 3, \tilde{a}^c \leq 1\} \sqrt{t} \varepsilon^{k-3}.$$

It is clear that for each fixed k , the right-hand side of (2.24) goes to zero as $\varepsilon \rightarrow 0$ since $a \geq 1$. In light of the convergence in (2.13), the proof of (2.5) is complete.

In the end we provide a uniform estimate in k when k is large, and from now on we fix the large constant C in (2.24). Considering the terms $\mathbf{1}\{k \leq 3\}t^{\frac{\tilde{a}^c}{2}-1}\varepsilon^a$ and $\mathbf{1}\{k > 3, \tilde{a}^c \leq 1\}\sqrt{t}\varepsilon^{k-3}$ from the above expression for $\mathcal{E}(k, a)$, we have

$$(2.25) \quad \begin{aligned} & \sum_{a=1}^k \binom{k+1}{a} \left(\mathbf{1}\{k \leq 3\}t^{\frac{\tilde{a}^c}{2}-1}\varepsilon^a + \mathbf{1}\{k > 3, \tilde{a}^c \leq 1\}\sqrt{t}\varepsilon^{k-3} \right) \\ & \leq \mathbf{1}\{k \leq 3\}6 \left(1 + \frac{1}{\sqrt{t}} + \frac{1}{t} \right) \varepsilon + \mathbf{1}\{k > 3\} \left[(k+1) + \frac{(k+1)k}{2} \right] \sqrt{t}\varepsilon^{k-3}. \end{aligned}$$

For $\mathbf{1}\{k > 3, \tilde{a}^c \geq 2\}e^{-\frac{\tilde{a}^c \log \tilde{a}^c}{C}}t^{\frac{\tilde{a}^c}{2}-1}\varepsilon^a$, the remaining term in $\mathcal{E}(k, a)$, we write

$$e^{-\frac{\tilde{a}^c \log \tilde{a}^c}{C}}t^{\frac{\tilde{a}^c}{2}-1}\varepsilon^a = \frac{1}{t} \left(\frac{\sqrt{t}}{(\tilde{a}^c)^{\frac{1}{C}}} \right)^{\tilde{a}^c} \varepsilon^a,$$

and we choose M sufficiently large (only depending on C, t) so that when $\tilde{a}^c > M$,

$$\frac{\sqrt{t}}{(\tilde{a}^c)^{\frac{1}{C}}} + \varepsilon < \frac{\sqrt{t}}{M^{\frac{1}{C}}} + \varepsilon < \frac{1}{C}.$$

Thus,

$$e^{-\frac{\tilde{a}^c \log \tilde{a}^c}{C}}t^{\frac{\tilde{a}^c}{2}-1}\varepsilon^a \leq \mathbf{1}\{\tilde{a}^c > M\} \frac{1}{t} \left(\frac{\sqrt{t}}{M^{\frac{1}{C}}} \right)^{\tilde{a}^c} \varepsilon^a + \mathbf{1}\{\tilde{a}^c \leq M\}(1+t^{\frac{M}{2}-1})\varepsilon^{k-M}.$$

Thus, when $k > M$, we have

$$(2.26) \quad \begin{aligned} & \sum_{a=1}^k \binom{k+1}{a} \mathbf{1}\{\tilde{a}^c \geq 2\} e^{-\frac{\tilde{a}^c \log \tilde{a}^c}{C}}t^{\frac{\tilde{a}^c}{2}-1}\varepsilon^a \\ & \leq \left[\sum_{a=1}^k \binom{k+1}{a} \frac{1}{t} \left(\frac{\sqrt{t}}{M^{\frac{1}{C}}} \right)^{\tilde{a}^c} \varepsilon^a + \sum_{a=1}^k \binom{k+1}{a} (1+t^{\frac{M}{2}-1})\varepsilon^{k-M} \right] \\ & \leq \left[\frac{M^{\frac{1}{C}}}{t^{\frac{3}{2}}} \left(\frac{\sqrt{t}}{M^{\frac{1}{C}}} + \varepsilon \right)^{k+1} + 2^{k+1}(1+t^{\frac{M}{2}-1})\varepsilon^{k-M} \right]. \end{aligned}$$

Combining (2.24), (2.25) and (2.26), we conclude that when $k > M$,

$$\begin{aligned} & \left\| \frac{1}{\varepsilon^{1+k}} e^{\frac{t}{\varepsilon} \mathcal{I}(v) + \frac{x}{\varepsilon} \mathcal{I}'(v)} \psi_{\varepsilon, k} \right\|_{L^2(\Delta_{k, \varepsilon}(t) \times \mathbb{R}^k)}^2 \\ & \leq C^k \varepsilon^{k-2} + C^k \left[(k+1) + \frac{(k+1)k}{2} \right] \sqrt{t}\varepsilon^{k-3} + C^k \left[\frac{M^{\frac{1}{C}}}{t^{\frac{3}{2}}} \left(\frac{\sqrt{t}}{M^{\frac{1}{C}}} + \varepsilon \right)^{k+1} + 2^{k+1}(1+t^{\frac{M}{2}-1})\varepsilon^{k-M} \right]. \end{aligned}$$

The right-hand side is summable in $k > M$ when $\varepsilon \ll 1$. Recalling the estimates in (2.14) and (2.15), we complete the proof of (2.6). \square

3. EXACTLY SOLVABLE BETA POLYMER: MOMENT CONVERGENCE

In this section, we study the limit of the exact moment formulas for the Beta polymer under the scaling of Theorem 1.3. Such moments formulas were first found in the work of [4], though the point-to-point moments we consider here are given explicitly in [9]. As $\varepsilon \rightarrow 0$, we prove on the formal level that these formulas converge to the corresponding SHE moment formulas given in [5, Section 6.2]. We will not provide herein a rigorous proof of these moment formula asymptotics (similar asymptotics are present, for instance, in [5]) but rather just work at the level of critical point analysis of the contour integrals. This provides an independent confirmation of the correctness of Theorem 1.3 (though even a rigorous proof of these moment formula convergences would not imply Theorem 1.3 due to the lack of well-posedness of the moment problem for the SHE).

Recall that for the Beta polymer, $B_{i,j}$ has $\text{Beta}(\alpha, \beta)$ distribution. Let $\mu = \alpha$ and $\nu = \alpha + \beta$, then it follows from [4, 9] that for $T \in \mathbb{Z}_{\geq 0}$ and $n_1 \geq \dots \geq n_k \in \mathbb{Z}$,

$$(3.1) \quad \mathbb{E}\{Z(T, n_1) \dots Z(T, n_k)\} = \frac{(\nu)_k}{(2\pi i)^k} \int \dots \int \prod_{1 \leq A < B \leq k} \frac{z_A - z_B}{z_A - z_B - 1} \prod_{j=1}^k \left(\frac{\nu + z_j}{z_j} \right)^{n_j} \left(\frac{\mu + z_j}{\nu + z_j} \right)^T \frac{dz_j}{(\nu + z_j)^2},$$

where the contour for z_k is a small circle around the origin, and the contour for z_A contains the contour for $z_B + 1$ for all $1 \leq A < B \leq k$, as well as the origin, but all contours exclude $-\nu$. Here $(\nu)_k = \nu(\nu + 1) \dots (\nu + k - 1)$.

There is a similar equation given in [5, Section 6.2] for the moments of the SHE. In terms of the scalings of Theorem 1.3, it implies that for $t \geq 0$ and $x_1 \geq \dots \geq x_k$,

$$(3.2) \quad \mathbb{E}\{\mathcal{V}(t, x_1) \dots \mathcal{V}(t, x_k)\} = \frac{(1 - 2\gamma)^{2k}}{(2\pi i)^k 2^k (1 - \gamma)^{2k}} \int \dots \int \prod_{1 \leq A < B \leq k} \frac{z_A - z_B}{z_A - z_B - 1} \prod_{j=1}^k \exp \left(\frac{(1 - 2\gamma)^4 t z_j^2}{8\gamma(1 - \gamma)} - \frac{(1 - 2\gamma)^2 x_j z_j}{2\gamma(1 - \gamma)} \right) dz_j,$$

where the contour for z_j is $r_j + i\mathbb{R}$ for arbitrary real r_1, \dots, r_k such that $r_j > r_{j+1} + 1$ for all j .

The goal is to prove the convergence of (3.1) to (3.2) after a proper rescaling, and as we stressed at the beginning, we will not provide a detailed rigorous proof but only sketch the critical point analysis. We fix $t > 0, x_j \in \mathbb{R}$ and $\gamma \in (0, \frac{1}{2})$, and let $\alpha = \beta = \varepsilon^{-1}$, and (recalling $(t_\varepsilon, x_\varepsilon)$ defined before Theorem 1.1) $T = \frac{t_\varepsilon}{\varepsilon^2}$, $n_j = \gamma \frac{t_\varepsilon}{\varepsilon^2} + \frac{x_{\varepsilon,j}}{\varepsilon}$. We define

$$f_\varepsilon(z, n) = \log \frac{\alpha + z}{2\alpha + z} + \frac{n}{T} \log \frac{2\alpha + z}{z}, \quad \text{so} \quad \left(\frac{2\alpha + z_j}{z_j} \right)^{n_j} \left(\frac{\alpha + z_j}{2\alpha + z_j} \right)^T = e^{T f_\varepsilon(z_j, n_j)}$$

A simple calculation shows that the critical point of f_ε is given by

$$(3.3) \quad \frac{2n}{T - 2n} \alpha \sim \frac{2\gamma}{(1 - 2\gamma)} \varepsilon^{-1} =: z_{0,\varepsilon}.$$

We may deform our contours to lie close to the critical point $z_{0,\varepsilon}$. Since $\varepsilon \ll 1$, in the vicinity of $z_{0,\varepsilon}$ the contours can be approximated by vertical straight lines of length on the order of ε^{-1} . We will assume (without proof) that with small error, we can replace the integrand by its approximation around the critical point (such an argument would involve describing steep-descent contours and similar examples can be found, for instance, in [5]). Taylor expanding to second order around $z_{0,\varepsilon}$ and setting $z_j = z_{0,\varepsilon} + \tilde{z}_j$, we find

$$f_\varepsilon(z_j, n_j) \approx -\mathcal{I}(1 - 2\gamma) + 2\mathcal{I}'(1 - 2\gamma) \frac{\varepsilon x_{\varepsilon,j}}{t_\varepsilon} + \frac{(1 - 2\gamma)^4 \varepsilon^2}{8\gamma(1 - \gamma)} \tilde{z}_j^2 - \frac{(1 - 2\gamma)^2 \varepsilon^2 x_{\varepsilon,j}}{2\gamma(1 - \gamma) t_\varepsilon} \tilde{z}_j + O(\varepsilon^3),$$

hence (under the aforementioned critical point hypothesis) we find that

$$\begin{aligned} & \mathbb{E} \left\{ \prod_{j=1}^k \frac{1}{\varepsilon} e^{\frac{t_\varepsilon}{\varepsilon^2} \mathcal{I}(1 - 2\gamma) - \frac{2x_{\varepsilon,j}}{\varepsilon} \mathcal{I}'(1 - 2\gamma)} Z \left(\frac{t_\varepsilon}{\varepsilon^2}, \gamma \frac{t_\varepsilon}{\varepsilon^2} + \frac{x_{\varepsilon,j}}{\varepsilon} \right) \right\} \\ & \rightarrow \frac{(1 - 2\gamma)^{2k}}{(2\pi i)^k 2^k (1 - \gamma)^{2k}} \int \dots \int \prod_{1 \leq A < B \leq k} \frac{\tilde{z}_A - \tilde{z}_B}{\tilde{z}_A - \tilde{z}_B - 1} \prod_{j=1}^k \exp \left(\frac{(1 - 2\gamma)^4 t \tilde{z}_j^2}{8\gamma(1 - \gamma)} - \frac{(1 - 2\gamma)^2 x_j \tilde{z}_j}{2\gamma(1 - \gamma)} \right) d\tilde{z}_j, \end{aligned}$$

where the contours are as in (3.2). The right-hand side equals $\mathbb{E}\{\mathcal{V}(t, x_1) \dots \mathcal{V}(t, x_k)\}$ as desired.

APPENDIX A. SHARP LARGE DEVIATION FOR SYMMETRIC SIMPLE RANDOM WALK

Recall that

$$\mathcal{P}_\varepsilon(t_\varepsilon, x_\varepsilon, m_1, m_2) = \mathbf{P}^0 \left(S_{\frac{t_\varepsilon}{\varepsilon^2} + m_1} = v \frac{t_\varepsilon}{\varepsilon^2} + \frac{x_\varepsilon}{\varepsilon} + m_2 \right),$$

$p_{\sigma^2}(t, x)$ is the density of $N(0, \sigma^2 t)$, and assume that $m_1, m_2 \in \mathbb{Z}$, $|m_1|, |m_2| \leq 1$ and $m_1 - m_2$ is even.

Lemma A.1. *For fixed $t > 0, x \in \mathbb{R}$, we have*

$$(A.1) \quad \frac{1}{\varepsilon} e^{\frac{t_\varepsilon}{\varepsilon^2} \mathcal{I}(v) + \frac{x_\varepsilon}{\varepsilon} \mathcal{I}'(v)} \mathcal{P}_\varepsilon(t_\varepsilon, x_\varepsilon, m_1, m_2) \rightarrow \frac{2p_{1-v^2}(t, x)}{(1+v)^{\frac{m_1+m_2}{2}} (1-v)^{\frac{m_1-m_2}{2}}}$$

as $\varepsilon \rightarrow 0$, and there exists a constant $C > 0$ such that for all $t > 0, x \in \mathbb{R}$,

$$(A.2) \quad \frac{1}{\varepsilon} e^{\frac{t_\varepsilon}{\varepsilon^2} \mathcal{I}(v) + \frac{x_\varepsilon}{\varepsilon} \mathcal{I}'(v)} \mathcal{P}_\varepsilon(t_\varepsilon, x_\varepsilon, m_1, m_2) \leq C \left(\frac{1}{\varepsilon} \mathbf{1}\{t \leq 100\varepsilon^2, |x| \leq C\varepsilon\} + \mathbf{1}\{t > 100\varepsilon^2\} \frac{1}{\sqrt{t}} e^{-\frac{x^2}{4t}} \right).$$

Proof. To simplify the notation, let

$$(A.3) \quad n = \frac{t_\varepsilon}{\varepsilon^2} + m_1, \quad m = v \frac{t_\varepsilon}{\varepsilon^2} + \frac{x_\varepsilon}{\varepsilon} + m_2,$$

so that $\mathcal{P}_\varepsilon(t_\varepsilon, x_\varepsilon, m_1, m_2) = \mathbf{P}^0(S_n = m)$. We will utilize the notation \lesssim when the left-hand side is bounded by a constant (independent of ε and t and x) times the right-hand side for all ε sufficiently small. We first prove the estimate (A.2) that is uniform in $t > 0, x \in \mathbb{R}$, then prove the convergence in (A.1) for fixed $t > 0, x \in \mathbb{R}$.

Case 1: $t \leq 100\varepsilon^2$. It is clear that $t_\varepsilon \leq 100\varepsilon^2$, and by the fact that $|m| \leq n$ (or else $\mathbf{P}^0(S_n = m) = 0$), there exists $C > 0$ so that $|x_\varepsilon| \leq C\varepsilon$. Thus, we have

$$\frac{1}{\varepsilon} e^{\frac{t_\varepsilon}{\varepsilon^2} \mathcal{I}(v) + \frac{x_\varepsilon}{\varepsilon} \mathcal{I}'(v)} \mathcal{P}_\varepsilon(t_\varepsilon, x_\varepsilon, m_1, m_2) \lesssim \frac{1}{\varepsilon} \mathbf{1}\{|x_\varepsilon| \leq C\varepsilon\}.$$

This proves the first bound on the right-hand side of (A.2).

Case 2: $t > 100\varepsilon^2$. Recall that $|m_1|, |m_2| \leq 1$, so $n \geq 90$ – we will implicitly use the largeness of n in some of the bounds below. By (A.3),

$$\frac{t_\varepsilon}{\varepsilon^2} \mathcal{I}(v) + \frac{x_\varepsilon}{\varepsilon} \mathcal{I}'(v) = \frac{1}{2} n \log(1-v^2) + \frac{1}{2} m \log \frac{1+v}{1-v} - m_1 \mathcal{I}(v) + (vm_1 - m_2) \mathcal{I}'(v),$$

which implies

$$e^{\frac{t_\varepsilon}{\varepsilon^2} \mathcal{I}(v) + \frac{x_\varepsilon}{\varepsilon} \mathcal{I}'(v)} \lesssim e^{\frac{1}{2} n \log(1-v^2) + \frac{1}{2} m \log \frac{1+v}{1-v}}.$$

Using Stirling's approximation,

$$\sqrt{2\pi} N^{N+\frac{1}{2}} e^{-N} \leq N! \leq e N^{N+\frac{1}{2}} e^{-N},$$

we derive the bound

$$\mathbf{P}^0(S_n = m) \lesssim \sqrt{\frac{n}{n^2 - m^2}} e^{-\frac{n+m}{2} \log(1+\frac{m}{n}) - \frac{n-m}{2} \log(1-\frac{m}{n})}$$

when $|m| < n$.

Now we consider three different subcases of case 2. Fix $0 < \tau \ll 1$, and let $k = m/n$.

Case 2.i: $|k| = 1$. We have $P(S_n = m) = 2^{-n}$, so

$$\frac{1}{\varepsilon} e^{\frac{t_\varepsilon}{\varepsilon^2} \mathcal{I}(v) + \frac{x_\varepsilon}{\varepsilon} \mathcal{I}'(v)} \mathcal{P}_\varepsilon(t_\varepsilon, x_\varepsilon, m_1, m_2) \lesssim \frac{1}{\varepsilon} e^{\frac{1}{2} n \log(1-v^2) + \frac{1}{2} m \log \frac{1+v}{1-v} - n \log 2}.$$

In both cases of $m = n$ and $m = -n$, using the fact that $v \in (0, 1)$, we have

$$\frac{1}{\varepsilon} e^{\frac{1}{2} n \log(1-v^2) + \frac{1}{2} m \log \frac{1+v}{1-v} - n \log 2} \leq \frac{1}{\varepsilon} e^{-\delta n}$$

for some $\delta > 0$ depending on v . Furthermore,

$$\frac{1}{\varepsilon} e^{-\delta n} = \frac{1}{\sqrt{\varepsilon^2 n}} \sqrt{n} e^{-\frac{\delta n}{2}} e^{-\frac{\delta n}{2}} \lesssim \frac{1}{\sqrt{t}} e^{-\frac{\delta n}{2}},$$

so it remains to show that $x^2/t \lesssim n$. By the fact that $|m| \leq n$, we have

$$-t_\varepsilon \lesssim \varepsilon x_\varepsilon \lesssim t_\varepsilon,$$

which implies $-t \lesssim \varepsilon x \lesssim t$. This implies that in this case, the second bound on the right-hand side of (A.2) holds.

Case 2.ii: $|k| < 1$ and $|k - v| \geq \tau$. We have

$$(A.4) \quad \frac{1}{\varepsilon} e^{\frac{t_\varepsilon}{\varepsilon^2} \mathcal{I}(v) + \frac{x_\varepsilon}{\varepsilon} \mathcal{I}'(v)} \mathcal{P}_\varepsilon(t_\varepsilon, x_\varepsilon, m_1, m_2) \lesssim \frac{1}{\varepsilon} \sqrt{\frac{n}{n^2 - m^2}} e^{-\frac{n}{2} F(k)},$$

where

$$(A.5) \quad F(k) = (1+k) \log \frac{1+k}{1+v} + (1-k) \log \frac{1-k}{1-v}, \quad k \in (-1, 1).$$

It is straightforward to check that $F(k)$ attains its minimum at v and $F''(k)$ is bounded from below by some positive constant. Since $|k-v| \geq \tau$, we have $F(k) \geq \delta \tau^2$ with $\delta = \frac{1}{2} \min_{k \in (-1, 1)} F''(k)$. In addition, since $|m| < n$, we have $-n+2 \leq m \leq n-2$, so

$$\frac{1}{\varepsilon} \sqrt{\frac{n}{n^2 - m^2}} \leq \frac{1}{\sqrt{\varepsilon^2 n}} \frac{1}{\sqrt{1 - \left(\frac{m}{n}\right)^2}} \lesssim \frac{1}{\sqrt{t}} \sqrt{n}.$$

Therefore, we have

$$\frac{1}{\varepsilon} e^{\frac{t\varepsilon}{2} \mathcal{I}(v) + \frac{x\varepsilon}{\varepsilon} \mathcal{I}'(v)} \mathcal{P}_\varepsilon(t_\varepsilon, x_\varepsilon, m_1, m_2) \lesssim \frac{1}{\sqrt{t}} \sqrt{n} e^{-\frac{n}{2} \delta \tau^2} \lesssim \frac{1}{\sqrt{t}} e^{-\frac{n}{4} \delta \tau^2},$$

and the same discussion as in case 2.i shows $e^{-\frac{n}{4} \delta \tau^2} \lesssim e^{-\frac{x^2}{Ct}}$ for some $C > 0$, matching the second bound on the right-hand side of (A.2).

Case 2.iii: $|k-v| < \tau$. We have

$$\frac{1}{\varepsilon} \sqrt{\frac{n}{n^2 - m^2}} \lesssim \frac{1}{\sqrt{t}}.$$

For the exponent, it is clear that $F(k) \geq \delta(k-v)^2$ with the same $\delta > 0$ from case 2.ii, so

$$e^{-\frac{n}{2} F(k)} \leq e^{-\frac{n}{2} \delta (k-v)^2}.$$

Since $k = m/n$ with m, n given in (A.3), we have

$$(A.6) \quad n(k-v)^2 = \frac{(x_\varepsilon + (m_2 - vm_1)\varepsilon)^2}{t_\varepsilon + m_1 \varepsilon^2},$$

so

$$e^{-\frac{n}{2} \delta (k-v)^2} \lesssim e^{-\frac{(|x| - \tilde{C}\varepsilon)^2}{Ct}} \leq e^{-\frac{x^2}{Ct}} e^{\frac{2\tilde{C}\varepsilon|x|}{Ct}}$$

for some $C, \tilde{C} > 0$. For $M > 0$, if $|x| \leq M\varepsilon$, we have the desired estimate; if $|x| > M\varepsilon$, we have

$$e^{-\frac{x^2}{Ct}} e^{\frac{2\tilde{C}\varepsilon|x|}{Ct}} < e^{-\frac{x^2}{Ct}} e^{\frac{2\tilde{C}M\varepsilon^2}{Ct}},$$

so we only need to choose M sufficiently large to complete the proof of (A.2).

To prove (A.1), we note that for fixed $t > 0, x \in \mathbb{R}$ and sufficiently small $\varepsilon, |k-v| \ll 1$ (we are in the region of case 2.iii). We use Stirling's approximation and the fact that $n, \frac{n+m}{2}, \frac{n-m}{2} \rightarrow \infty$ to obtain that

$$\frac{\mathbf{P}^0(S_n = m)}{\sqrt{\frac{2n}{\pi(n^2 - m^2)}} e^{-\frac{n}{2} [(1+k) \log(1+k) + (1-k) \log(1-k)]}} \rightarrow 1$$

as $\varepsilon \rightarrow 0$. Thus we only need to analyze

$$\begin{aligned} & \frac{1}{\varepsilon} e^{\frac{t\varepsilon}{2} \mathcal{I}(v) + \frac{x\varepsilon}{\varepsilon} \mathcal{I}'(v)} \sqrt{\frac{2n}{\pi(n^2 - m^2)}} e^{-\frac{n}{2} [(1+k) \log(1+k) + (1-k) \log(1-k)]} \\ &= \sqrt{\frac{2n}{\pi(n^2 - m^2)\varepsilon^2}} \frac{e^{-\frac{n}{2} F(k)}}{(1+v)^{\frac{m_1+m_2}{2}} (1-v)^{\frac{m_1-m_2}{2}}}, \end{aligned}$$

with $F(k)$ defined in (A.5). First, we have

$$\sqrt{\frac{2n}{\pi(n^2 - m^2)\varepsilon^2}} \rightarrow \sqrt{\frac{2}{\pi(1-v^2)t}}.$$

Secondly, by (A.6), we have $|k-v| \lesssim \varepsilon$ for fixed $t > 0, x \in \mathbb{R}$ and $n(k-v)^2 \rightarrow \frac{x^2}{t}$ as $\varepsilon \rightarrow 0$. We expand $F(k) = \frac{1}{2} F''(v)(k-v)^2 + O(|k-v|^3)$ and conclude that

$$\frac{n}{2} F(k) \rightarrow F''(v) \frac{x^2}{4t} = \frac{x^2}{2(1-v^2)t}$$

as $\varepsilon \rightarrow 0$. This completes the proof of (A.1). \square

REFERENCES

- [1] T. Alberts, K. Khanin, J. Quastel. The intermediate disorder regime for directed polymers in dimension $1 + 1$. *Ann. Probab.*, 42:1212–1256, 2014.
- [2] G. Amir, I. Corwin, J. Quastel. Probability distribution of the free energy of the continuum directed random polymer in $1 + 1$ dimensions. *Commun. Pure Appl. Math.*, **64**:466–537, 2011.
- [3] M. Balázs, F. Rassoul-Agha, T. Seppäläinen. The random average process and random walk in a space-time random environment in one dimension. *Commun. Math. Phys.*, 266:499–545, 2006.
- [4] G. Barraquand and I. Corwin. Random-walk in beta-distributed random environment. *Probab. Theo. Rel. Fields*, onlinefirst, 2016.
- [5] A. Borodin, I. Corwin. Macdonald processes. *Probab. Theo. Rel. Fields*, **158**:225–400, 2014.
- [6] F. Caravenna, R. Sun, and N. Zygouras. Polynomial chaos and scaling limits of disordered systems. *J. Eur. Math. Soc.*, to appear.
- [7] I. Corwin. The Kardar-Parisi-Zhang equation and universality class. *Rand. Mat.: Theo. Appl.*, **1**:1130001, 2012.
- [8] I. Corwin. Macdonald processes, quantum integrable systems and the Kardar-Parisi-Zhang universality class. *Proceedings of the International Congress of Mathematicians 2014*.
- [9] T. Thiery, P. Le Doussal. Exact solution for a random walk in a time-dependent 1D random environment: the point-to-point Beta polymer *arXiv preprint arXiv:1605.07538*, 2016.
- [10] F. Rassoul-Agha, T. Seppäläinen, A. Yilmaz. Quenched free energy and large deviations for random walks in random potentials. *Commun. Pure Appl. Math.*, **66**:202–244, 2013.

I. CORWIN, COLUMBIA UNIVERSITY, DEPARTMENT OF MATHEMATICS, 2990 BROADWAY, NEW YORK, NY 10027, USA, AND CLAY MATHEMATICS INSTITUTE, 10 MEMORIAL BLVD. SUITE 902, PROVIDENCE, RI 02903, USA
E-mail address: `ivan.corwin@gmail.com`

Y. GU, DEPARTMENT OF MATHEMATICS, BUILDING 380, STANFORD UNIVERSITY, STANFORD, CA, 94305, USA
E-mail address: `yg@math.stanford.edu`